Journal of Computational and Applied Mathematics 30 (1990) 87–97
North-Holland

An existence and nonexistence theorem for solutions of nonlinear systems and its application to algebraic equations

Tetsuro YAMAMOTO and Xiaojun CHEN
Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan

Received 30 April 1989

Abstract: In this paper, we first present a componentwise existence and nonexistence theorem for solutions of nonlinear equations in $\mathbb{R}^n$ or $\mathbb{C}^n$, which improves those of Alefeld (1986) and Yamamoto (1978,1979,1980,1984). Next, the theorem is applied to algebraic equations. Finally, the results are illustrated with examples.

Keywords: Existence and nonexistence theorem, nonlinear equations, a posteriori componentwise error estimates, algebraic equations.

1. Introduction

There are many existence theorems for solutions of nonlinear equations which are applicable to componentwise error estimates of approximate solutions obtained by some methods (cf. [6,7,9] etc.). Among others, Yamamoto [10–14] obtained some related results and recently Alefeld [2] generalized a result of [14].

In this paper, in Section 3, after the preliminary section (Section 2), we shall give an existence and nonexistence theorem for a solution of an equation under Kantorovich-type assumptions which are weaker than those of [2] and [10–12,14]. The theorem seems to be new and sharper.

Next, in Section 4, we shall apply our results to algebraic polynomials to obtain a Gerschgorin-type existence theorem of solutions under computationally verifiable conditions. The assumptions are stronger than Smith's theorem [8]. However, our result guarantees existence of a solution in each of n disks D_i, $i=1, 2, \ldots, n$, while Smith's theorem only asserts that any connected component of the union of n circular regions Γ_i, consisting of just m disks, contains exactly m zeros. Our results improve those of Zheng [15] too.

Finally, in Section 5, our results will be illustrated with numerical examples.

2. Preliminaries

Throughout this paper, according to [7,9,10,14], we use the following notation and definitions. Let $x=(x_i)$, $y=(y_i)\in\mathbb{R}^n$, $A=(a_{ij})$, $B=(b_{ij})\in\mathbb{R}^{n\times n}$, $H=(h_{ijk})\in\mathbb{R}^{n\times n\times n}$. We define

$$\nu[x]=(|x_i|),\qquad \rho(x, y)=\nu[x-y],$$
$$\nu[A]=(|a_{ij}|),\qquad \rho(A, B)=\nu[A-B],\qquad \nu[H]=(|h_{ijk}|).$$

We write

$$\begin{array}{llll} x \geqslant y & \text{or} \quad y \leqslant x & \text{if } x_i \geqslant y_i, & i = 1, 2, \ldots, n, \\ A \geqslant B & \text{or} \quad B \leqslant A & \text{if } a_{ij} \geqslant b_{ij}, & i, j = 1, 2, \ldots, n, \\ H \geqslant 0 & \text{or} \quad 0 \leqslant H & \text{if } h_{ijk} \geqslant 0, & i, j, k = 1, 2, \ldots, n. \end{array}$$

For a nonnegative vector $v \geqslant 0$, we put

$$U(x^{(0)}, v) = \{ x \in \mathbb{R}^n \mid \rho(x^{(0)}, x) \leqslant v \}.$$

In [12], Yamamoto derived the following relations for a vector $u \geqslant 0$, a matrix $K = (\kappa_{ij}) \geqslant 0$ and a third-order tensor $H = (h_{ijk}) \geqslant 0$:

$$Ku \leqslant \| u \|_\infty \, \kappa_1, \qquad Hu^2 \leqslant \| u \|_\infty^2 h_1,$$

and

$$Ku \leqslant \| u \|_1 \, \kappa_\infty, \qquad Hu^2 \leqslant \| u \|_1^2 h_\infty,$$

where

$$\| u \|_\infty = \max_i u_i, \qquad \kappa_1 = \left(\sum_{j=1}^n \kappa_{ij} \right) \in \mathbb{R}^n, \qquad h_1 = \left(\sum_{j=1}^n \sum_{k=1}^n h_{ijk} \right) \in \mathbb{R}^n,$$

$$\| u \|_1 = \sum_{i=1}^n u_i, \qquad \kappa_\infty = \left(\max_j \kappa_{ij} \right) \in \mathbb{R}^n, \qquad h_\infty = \left(\max_{j,k} h_{ijk} \right) \in \mathbb{R}^n.$$

Alefeld generalized those as follows.

Lemma 1. *Let $u = (u_i) \in \mathbb{R}^n$, $u \geqslant 0$, $K = (\kappa_{ij}) \in \mathbb{R}^{n \times n}$, $K \geqslant 0$, $H = (h_{ijk}) \in \mathbb{R}^{n \times n \times n}$ and $H \geqslant 0$. Then it holds for $p \geqslant 1$, $q \geqslant 1$, $p^{-1} + q^{-1} = 1$, that*

$$Ku \leqslant \| u \|_p \, \kappa_q, \qquad Hu^2 \leqslant \| u \|_p^2 h_q, \tag{2.1}$$

where

$$\| u \|_p = \left(\sum_{i=1}^n u_i^p \right)^{1/p}, \qquad \kappa_q = \left(\left(\sum_{j=1}^n \kappa_{ij}^q \right)^{1/q} \right) \in \mathbb{R}^n$$

and

$$h_q = \left(\left(\sum_{j=1}^n \sum_{k=1}^n h_{ijk}^q \right)^{1/q} \right) \in \mathbb{R}^n.$$

In the inequalities (2.1), we understand that $q = \infty$ if $p = 1$ and $q = 1$ if $p = \infty$.

Remark 2. From Lemma 1, we have

$$\| K \|_p = \max_{\| x \|_p = 1} \| Kx \|_p \leqslant \| \kappa_q \|_p$$

and

$$\| H \|_p = \max_{\| x \|_p = \| y \|_p = 1} \| Hxy \|_p \leqslant \| h_q \|_p.$$

Furthermore, we have

$$\|K\|_1 = \max_j \sum_{i=1}^n \kappa_{ij} \leqslant \|\kappa_\infty\|_1, \qquad \|H\|_1 = \max_{j,k} \sum_{i=1}^n h_{ijk} \leqslant \|h_\infty\|_1,$$

$$\|K\|_\infty = \max_i \sum_{j=1}^n \kappa_{ij} = \|\kappa_1\|_\infty, \qquad \|H\|_\infty \leqslant \max_i \sum_{j=1}^n \sum_{k=1}^n h_{ijk} \leqslant \|h_1\|_\infty.$$

3. An existence and nonexistence theorem for a solution of an equation

We consider the nonlinear equation

$$f(x) = 0, \quad x \in D \subset \mathbb{R}^n, \tag{3.1}$$

where f is Fréchet-differentiable. We assume that A is a nonsingular matrix which approximates $f'(x^{(0)})$, $H \geqslant 0$ is a third-order tensor, and for any x, $y \in D$ inequalities

$$\rho(A^{-1}f'(x), A^{-1}f'(y)) \leqslant H\rho(x, y)$$

and

$$\|A^{-1}(f'(x) - f'(y))\|_p \leqslant \gamma \|x - y\|_p$$

hold, where γ is a nonnegative constant with $\gamma \leqslant \|H\|_p$. We put

$$K = \nu[A^{-1}(f'(x^{(0)}) - A)], \qquad \epsilon = \nu[A^{-1}f(x^{(0)})] = (\epsilon_i),$$
$$m = \|K\|_p, \qquad \eta = \|\epsilon\|_p.$$

Furthermore, without loss of generality we assume that $\epsilon \neq 0$, which means that $x^{(0)}$ is not a solution.

Theorem 3. *Under the above notation and definitions, we have the following*:

(i) *If*

$$m < 1, \quad a = (1 - m)^2 - 2\gamma\eta \geqslant 0, \tag{3.2}$$

and $S(x^{(0)}, t^*) = \{x \in \mathbb{R}^n \mid \|x^{(0)} - x\|_p \leqslant t^*\} \subset D$, *where* $t^* = 2\eta/(1 - m + \sqrt{a})$, *then there exists a solution* x^* *of* (3.1) *in* $U(x^{(0)}, u)$, *where*

$$u = \epsilon + t^*\kappa_q + \tfrac{1}{2}(t^*)^2 h_q.$$

(ii) *There is no solution in* $\operatorname{int}\{U(x^{(0)}, v) \cap D\}$, *where*

$$v = \frac{2\|\epsilon\|_\infty}{1 + \|K\|_\infty + \sqrt{(1 + \|K\|_\infty)^2 + 2\|H\|_\infty \|\epsilon\|_\infty}}(1, 1, \ldots, 1)^t.$$

Proof. (i) Since $\|K\|_p = \|A^{-1}(A - f'(x^{(0)}))\|_p = m < 1$, we have that $f'(x^{(0)})^{-1}$ exists;

$$\|f'(x^{(0)})^{-1}A\|_p \leqslant (1 - m)^{-1},$$

$$\|f'(x^{(0)})^{-1}(f'(x) - f'(y))\|_p \leqslant (1 - m)^{-1}\gamma\|x - y\|_p$$

and

$$\| f'(x^{(0)})^{-1} f(x^{(0)}) \|_p \leqslant (1-m)^{-1}\eta,$$

so that Kantorovich's condition $2(1-m)^{-2}\gamma\eta \leqslant 1$ for Newton's method applied to (3.1) is satisfied. Hence, according to the Kantorovich theorem, there exists a solution $x^* \in S(x^{(0)}, t^*)$, which is unique in

$$\tilde{S} = \begin{cases} S(x^{(0)}, t^{**}) & \text{if } a = 0, \\ \operatorname{int}\{ S(x^{(0)}, t^{**}) \cap D \} & \text{if } a > 0. \end{cases}$$

Then we obtain

$$\begin{aligned} \rho(x^{(0)}, x^*) &= \nu\big[A^{-1}f(x^{(0)}) + A^{-1}(A - f'(x^{(0)}))(x^{(0)} - x^*) \\ &\qquad + A^{-1}(f'(x^{(0)})(x^{(0)} - x^*) - f(x^{(0)}) + f(x^*))\big] \\ &\leqslant \epsilon + \nu\big[A^{-1}(A - f'(x^{(0)}))\big]\nu\big[x^{(0)} - x^*\big] \\ &\qquad + \nu\left[A^{-1}\left(f'(x^{(0)}) - \int_0^1 f'(x^* + t(x^{(0)} - x^*))\,\mathrm{d}t\right)\right]\nu\big[x^{(0)} - x^*\big] \\ &\leqslant \epsilon + \| x^{(0)} - x^* \|_p\, \kappa_q + \tfrac{1}{2}\| x^{(0)} - x^* \|_p^2 h_q \\ &\leqslant \epsilon + t^*\kappa_q + \tfrac{1}{2}(t^*)^2 h_q = u, \end{aligned}$$

that is, $x^* \in U(x^{(0)}, u)$.

(ii) If there is a solution $y^* \in \operatorname{int}\{U(x^{(0)}, v) \cap D\}$, then

$$\begin{aligned} A^{-1}f(x^{(0)}) &= A^{-1}(f(x^{(0)}) - f(y^*)) \\ &= A^{-1}\left[\int_0^1 \{f'(y^* + t(x^{(0)} - y^*)) - f'(x^{(0)})\}(x^{(0)} - y^*)\,\mathrm{d}t \right. \\ &\qquad \left. + (f'(x^{(0)}) - A)(x^{(0)} - y^*)\right] + (x^{(0)} - y^*), \end{aligned}$$

so that

$$\begin{aligned} \epsilon = \nu\big[A^{-1}f(x^{(0)})\big] &\leqslant \tfrac{1}{2}H\nu\big[x^{(0)} - y^*\big]^2 + K\nu\big[x^{(0)} - y^*\big] + \nu\big[x^{(0)} - y^*\big] \\ &< \tfrac{1}{2}Hv^2 + Kv + v. \end{aligned}$$

This implies $\| \epsilon \|_\infty < \frac{1}{2}\| H \|_\infty \| v \|_\infty^2 + (\| K \|_\infty + 1)\| v \|_\infty = \| \epsilon \|_\infty$, which is a contradiction. □

Remark 4.. The results of [2] and [10,14] have been derived under stronger conditions

$$\| \kappa_q \|_p < 1 \quad \text{and} \quad \big(1 - \| \kappa_q \|_p\big)^2 - 2\| h_q \|_p \| \epsilon \|_p \geqslant 0 \tag{3.3}$$

than (3.2). Therefore it follows from Remark 2 that our existence domain improves the ones obtained in [2,10,14].

Remark 5. The nonexistence domain of Theorem 3 improves Alefeld's one. In fact, he assumed that $\delta = \min_i \epsilon_i > 0$ and proved that (3.1) has no solution in $U(x^{(0)},\beta)$, where

$$\beta = \frac{2\delta}{1 + \| \kappa_1 \|_\infty + \sqrt{(1 + \| \kappa_1 \|_\infty)^2 + 2\| h_1 \|_\infty \delta}}(1, 1, \ldots, 1)^{\mathrm{t}}.$$

Clearly we have $\beta \leqslant v$.

4. Application to algebraic equations

In this section, we consider the nth-degree polynomial

$$P(z)=z^n+a_1z^{n-1}+\cdots+a_n=\prod_{i=1}^{n}(z-z_i^*), \tag{4.1}$$

with complex coefficients a_i. To find all zeros z_i^*, $i=1, 2, \ldots, n$, of the polynomial, Durand [3] and Kerner [5] considered a simultaneous iteration process

$$x_i^{(k)}=x_i^{(k-1)}-\frac{P\left(x_i^{(k-1)}\right)}{\prod_{j\neq i}^{n}\left(x_i^{(k-1)}-x_j^{(k-1)}\right)}, \quad 1\leqslant i\leqslant n, \quad k=1, 2, \ldots, \tag{4.2}$$

provided that z_i^* are distinct, where $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\in\mathbb{C}$ and $x_i^{(0)}\neq x_j^{(0)}$, $i\neq j$. This process is called Durand–Kerner's method or the D–K method, although Weierstrass also considered this.

In [5], Kerner showed that the D–K method is equivalent to Newton's method applied to the equation

$$f(x)=(f_1(x), f_2(x), \ldots, f_n(x))^{\mathrm{t}}=0, \quad x\in\mathbb{C}^n,$$

where

$$f_i(x)=b_i(x)-a_i, \quad b_i(x)=(-1)^i\sum_{1\leqslant j_1<\cdots<j_i\leqslant n}x_{j_1}\cdots x_{j_i}, \qquad i=1, 2, \ldots, n.$$

In [10], Yamamoto obtained n circular regions, each of which contains at least one zero of the polynomial.

In this section, as an application of Theorem 3, we shall derive new existence circular regions under a Kantorovich-type condition and nonexistence circular regions with no restriction. The radius of the former regions are smaller than those of Smith and Yamamoto.

We put

$$\beta_{ij}=|x_i^{(0)}-x_j^{(0)}|^{-1}, \quad \beta=\max_{i,j}\beta_{ij},$$

$$\epsilon_i=\frac{|P(x_i^{(0)})|}{\prod_{j\neq i}^{n}\beta_{ij}}, \quad \eta=\sum_{i=1}^{n}\epsilon_i,$$

$$\mu=\eta\beta, \quad b=1+\frac{2\mu}{n-2}.$$

Let $H=(h_{ijk})$ be defined as follows:

$$h_{ijk}=\begin{cases}\beta b^{n-2} & \text{if } j\neq k,\ i=j \text{ or } j\neq k,\ i=k,\\ 2\eta\beta^2b^{n-3} & \text{if } j\neq k,\ i\neq j,\ i\neq k,\\ 0 & \text{if } j=k.\end{cases}$$

We put

$$\gamma=\begin{cases}2\mu\beta b^{n-3}+2\beta b^{n-2} & \text{if } n\geqslant 3,\\ 2\beta & \text{if } n=2,\end{cases}$$

and $h=\gamma\eta$. Then we have the following theorem.

Theorem 6. (i) *If $h \leqslant \frac{1}{2}$, then we have the following*:

(a) *In each closed disk*

$$D_i\colon \left| z - x_i^{(0)} \right| \leqslant d_i = \epsilon_i + \tfrac{1}{2}(t^*)^2 h_0,$$

there exists at least one zero of $P(z)$, where

$$t^* = \frac{2\eta}{1+\sqrt{1-2\gamma\eta}} \quad \text{and} \quad h_0 = \max(\beta b^{n-2}, 2n\beta^2 b^{n-3}).$$

(b) *Any connected component of the union of the set D_i consisting of just m disks contains exactly m zeros of $P(z)$.*

(ii) *If $h < \frac{1}{2}$, then all the zeros of $P(z)$ are simple.*

Proof. (i) By Theorem 3 with $p = 1$, it is sufficient to prove that

$$\rho\left(f'(x^{(0)})^{-1}f'(x),\ f'(x^{(0)})^{-1}f'(y)\right) \leqslant H\rho(x, y) \tag{4.3}$$

and

$$\| f'(x^{(0)})^{-1}(f'(x) - f'(y)) \|_1 \leqslant \gamma \| x - y \|_1$$

for any $x,\ y \in S(x^{(0)}, 2\eta) = \{ x \mid \| x^{(0)} - x \|_1 \leqslant 2\eta \}$. From the identity

$$z^n + b_1(x)z^{n-1} + \cdots + b_n(x) = \prod_{i=1}^{n} (z - x_i),$$

we have

$$\sum_{i=1}^{n} \frac{\partial b_i(x)}{\partial x_j} z^{n-i} = -\prod_{\substack{i=1\\ i\neq j}}^{n} (z - x_i).$$

As was shown by Kerner [5], the ith row of $f'(x^{(0)})^{-1}$ is given by

$$\frac{-1}{\prod_{j\neq i}^{n}\left(x_i^{(0)} - x_j^{(0)}\right)}\left(x_i^{(0)\,n-1},\ x_i^{(0)\,n-2},\ \ldots,\ 1\right).$$

Therefore, putting $Q(z) = (z - x_1^{(0)}) \cdots (z - x_n^{(0)})$, we have

$$\begin{aligned}
&\nu\left[f'(x^{(0)})^{-1}(f'(x) - f'(y))\right]_{ij}\\
&\quad = \nu\left[\sum_{k=1}^{n} \frac{x_i^{(0)\,n-k}}{Q'\left(x_i^{(0)}\right)}\left(\frac{\partial b_k(x)}{\partial x_j} - \frac{\partial b_k(y)}{\partial y_j}\right)\right]\\
&\quad = \nu\left[\frac{1}{Q'\left(x_i^{(0)}\right)}\left(\prod_{l\neq j}^{n}\left(x_i^{(0)} - y_l\right) - \prod_{l\neq j}^{n}\left(x_i^{(0)} - x_l\right)\right)\right]\\
&\quad = \nu\left[\frac{1}{Q'\left(x_i^{(0)}\right)} \sum_{\substack{k=1\\ k\neq j}}^{n} \prod_{\substack{l\neq j\\ l\neq k}}^{n}\left(x_i^{(0)} - x_l - t_{ij}(y_l - x_l)\right)(x_k - y_k)\right]\\
&\quad \leqslant (\tilde{H}\nu[x - y])_{ij},
\end{aligned}$$

where $0 \leqslant t_{ij} \leqslant 1$ and $\tilde{H} = (\tilde{h}_{ijk})$ is defined by

$$\tilde{h}_{ijk} = \begin{cases} \left| \dfrac{1}{Q'\left(x_i^{(0)}\right)} \displaystyle\prod_{l \neq j,k}^{n} \left(x_i^{(0)} - \hat{x}_l\right) \right| & \text{if } j \neq k, \\ 0 & \text{if } j = k, \end{cases}$$

with $\hat{x}_l = x_l + t_{ij}(y_l - x_l)$. Furthermore, we can prove that for any $x,\ y \in S(x^{(0)}, 2\eta)$,

$$\tilde{h}_{ijk} \leqslant h_{ijk} \quad \text{and} \quad \| \tilde{H} \|_1 \leqslant \gamma. \tag{4.4}$$

In fact, we have for $j \neq k,\ i = j$,

$$\begin{aligned} \tilde{h}_{ijk} &= \left| \frac{1}{Q'\left(x_j^{(0)}\right)} \prod_{l \neq j,k} \left(x_j^{(0)} - \hat{x}_l\right) \right| \\ &\leqslant \left| \frac{\prod_{l \neq j,k} \left(x_j^{(0)} - x_l^{(0)}\right)}{Q'\left(x_j^{(0)}\right)} \prod_{l \neq j,k} \left(1 + \frac{x_l^{(0)} - \hat{x}_l}{x_j^{(0)} - x_l^{(0)}}\right) \right| \\ &\leqslant \frac{1}{\left|x_j^{(0)} - x_k^{(0)}\right|} \prod_{l \neq j,k} \left(1 + \beta \left|x_l^{(0)} - \hat{x}_l\right|\right) \\ &\leqslant \beta \left(1 + \frac{2\beta\eta}{n-2}\right)^{n-2} = h_{ijk}, \end{aligned}$$

where we used the fact that $b_1 = \cdots = b_m = c/m$ maximize the function

$$\prod_{i=1}^{m} (1 + b_i)$$

subject to the conditions

$$b_i \geqslant 0 \quad \text{and} \quad \sum_{i=1}^{m} b_i \leqslant c.$$

The same argument works for $j \neq k,\ i = k$. Furthermore, for $i \neq j,\ j \neq k,\ i \neq k$, we have

$$\begin{aligned} \tilde{h}_{ijk} &= \left| \frac{1}{Q'\left(x_i^{(0)}\right)} \prod_{l \neq j,k} \left(x_i^{(0)} - \hat{x}_l\right) \right| \\ &= \left| \frac{x_i^{(0)} - \hat{x}_i}{Q'\left(x_i^{(0)}\right)} \prod_{l \neq i,j,k} \left(x_i^{(0)} - x_l^{(0)} + x_l^{(0)} - \hat{x}_l\right) \right| \\ &= \left| \frac{x_i^{(0)} - \hat{x}_i}{\left(x_i^{(0)} - x_j^{(0)}\right)\left(x_i^{(0)} - x_k^{(0)}\right)} \prod_{l \neq i,j,k} \left[1 + \frac{x_l^{(0)} - \hat{x}_l}{x_i^{(0)} - x_l^{(0)}}\right] \right| \\ &\leqslant 2\eta\beta^2 \left(1 + \frac{2\beta\eta}{n-2}\right)^{n-3} = h_{ijk}. \end{aligned}$$

The final bound is obtained from the fact that under the conditions $b_i \geqslant 0$, $1 \leqslant i \leqslant n$, and $b_1 + \cdots + b_m \leqslant c$, the function

$$b_1 \prod_{i=2}^{m} (1 + b_i)$$

is maximized only if $b_2 = \cdots = b_m$ $(=(c-1)/m$ if $c \geqslant 1$, and $=0$ if $c \leqslant 1)$.

Furthermore, we have

$$\| \tilde{H} \|_1 = \max_{j,k} \sum_{i=1}^{n} \tilde{h}_{ijk} \leqslant \gamma = \begin{cases} 2\eta\beta^2\left(1 + \dfrac{2\beta\eta}{n-2}\right)^{n-3} + 2\beta\left(1 + \dfrac{2\beta\eta}{n-2}\right)^{n-2} & \text{if} \quad n \geqslant 3, \\ 2\beta & \text{if} \quad n = 2, \end{cases}$$

where we again used the fact that $b_1 = b_2 = \cdots = b_m = c/m$ maximize the function

$$\sum_{i=1}^{m} b_i \prod_{j \neq i}^{m} (1 + b_j)$$

subject to the conditions

$$b_i \geqslant 0 \quad \text{and} \quad \sum_{i=1}^{m} b_i \leqslant c.$$

This implies that if $h \leqslant \frac{1}{2}$, then, in each closed disk D_i there exists at least one zero of $P(z)$, and any connected component of the union of the set D_i consisting of just m disks contains at least m zeros of $P(z)$. This proves (i), since the total number of zeros of $P(z)$ is equal to n.

(ii) If $h < \frac{1}{2}$, then we can prove that $f'(z^*)^{-1}$ exists and $z_1^*, \ldots, z_n^*$ are distinct. □

Remark 7. In [15], Zheng obtained the error estimates

$$\| x^{(m)} - z^* \|_\infty \leqslant \frac{(1-2s)^m g^{2^m - 1}}{1 - (1-2s) g^{2^m}} \| \epsilon \|_\infty, \quad s = \| \epsilon \|_\infty \beta, \ m \geqslant 0,$$

for the D–K method under the conditions $s < \frac{1}{2}$ and

$$g = g(n, s) = (n-1) \frac{s}{1-s} \left(1 + \frac{s}{1-2s}\right)^{n-1} \frac{1}{1-2s} \leqslant 1.$$

Clearly, for fixed $s > 0$, we have $g(n,s) \to \infty$ at $n \to \infty$. That is, the quantity s such that $g \leqslant 1$ should be very small if n is large. On the contrary, the quantity h in Theorem 6 need not be small for large n, since

$$h \leqslant 2\mu^2 e^{2\mu} + 2\mu e^{2\mu} = 2\mu e^{2\mu}(1 + \mu),$$

and the positive root μ^* of the equation $2\mu e^{2\mu}$ $(1 + \mu) = \frac{1}{2}$ is in the open interval $(\frac{1}{8}, \frac{1}{7})$.

5. Numerical examples

In this section, we shall illustrate our results with some examples.

Example 8 (Yamamoto [12], Alefeld [2]). We first consider the algebraic eigenvalue problem

$$Ty = \lambda y, \tag{5.1}$$

where

$$T=\begin{pmatrix}1 & 1 & \frac{1}{2}\\ 1 & 1 & \frac{1}{4}\\ \frac{1}{2} & \frac{1}{4} & 2\end{pmatrix}$$

and λ and y are eigenvalues and eigenvectors of T, respectively. We assume that $y=(y_i)\in\mathbb{R}^n$ satisfies $\| y\|_2^2=\sum_{i=1}^n | y_i|^2=1$. Let $x=(y_1, \ldots, y_n, \lambda)^t$. Then (5.1) can be written in the form

$$f(x)=\begin{pmatrix}(T-\lambda I)y\\ \frac{1}{2}(1-\| y\|_2^2)\end{pmatrix}=0.$$

Furthermore, we have

$$f'(x)=\begin{pmatrix}T-\lambda I & -y\\ -y^t & 0\end{pmatrix},$$

$$f''(x)=\left(\begin{array}{cccc|c|cccc|ccc} 0 & \cdots & 0 & -1 & & 0 & \cdots & 0 & 0 & -1 & & \\ & & & & \cdots & & & & -1 & & \ddots & -1 \\ -1 & 0 & \cdots & 0 & & 0 & \cdots & -1 & 0 & & & 0\end{array}\right).$$

We choose an initial approximation $x^{(0)}=(-0.7, 0.7, 0.15, 0.0)^t$ and put $A=f'(x^{(0)})$. Then we have

$$K=0, \qquad H=\nu\left[f'(x^{(0)})^{-1}f''(x^{(0)})\right], \qquad \epsilon=\nu\left[f'(x^{(0)})^{-1}f(x^{(0)})\right].$$

Furthermore, we choose $p=1$, $q=\infty$. Then we obtain $\| H\|_1\|\epsilon\|_1=0.4074514<\frac{1}{2}$. We can thus conclude that there exists a solution x^* in $U(x^{(0)},u)$, where

$$u=(0.0475545, 0.0293976, 0.1691796, 0.0271919)^t.$$

It is interesting to remark that $\| h_1\|_\infty\|\epsilon\|_\infty=0.7518750$ and $\| h_\infty\|_1\|\epsilon\|_1=0.5163171$, so that the conditions of [2, Corollary 3.2] and [14, Theorem 1] are not satisfied. Consequently existence of a solution cannot be guaranteed by results of [2] and [14].

Example 9 (Aberth [1], Yamamoto [10]). To illustrate Theorem 6, we consider the polynomial

$$\begin{aligned}P(z)&=z^5-10z^4+43z^3-104z^2+150z-100\\ &=(z-2)(z^2-2z+5)(z^2-6z+10).\end{aligned}$$

To find the zeros of $P(z)$, we use the D–K method with the initial value

$$x_i^{(0)}=-\frac{a_1}{n}+r_0\exp\left[\left(\frac{2(i-1)\pi}{n}+\frac{\pi}{2n}\right)\sqrt{-1}\right], \quad i=1, 2, \ldots, n,$$

where $a_1=-10$, $n=5$ and $r_0=6$, which are chosen according to Aberth [1]. After 11 steps of iteration, we obtained approximate zeros $x_i^{(11)}$ of $P(z)$. Considering $x_i^{(11)}$ as $x_i^{(0)}$ in Theorem 6, we obtain $h=0.00273842707184<\frac{1}{2}$.

Table 1

i	$\mathrm{Re}(x_i^{(11)})$	$\mathrm{Im}(x_i^{(11)})$	r_i	d_i
1	3.00048377555635	0.99919417201716	0.00470031245654	0.00094138455511
2	0.99989737424306	2.00000976307389	0.00051545540778	0.00010441314536
3	1.99965263876727	0.00068117182684	0.00382394244572	0.00076611055295
4	1.00000057774750	−2.00000003400358	0.00000289343786	0.00000190075138
5	2.99996563368582	−0.99988507291431	0.00059969922130	0.00012126190806

The radii d_i of the disks D_i and the radii r_i of Smith's circular regions

$$\Gamma_i\colon\ |z - x_i^{(11)}| \leqslant r_i = n\left|\frac{P\left(x_i^{(11)}\right)}{\prod_{j\neq i}^{n}\left(x_i^{(11)} - x_j^{(11)}\right)}\right|,\quad 1 \leqslant i \leqslant n,$$

are also shown in Table 1 together with $x_i^{(11)}$, $i = 1, 2, \dots, n$.

We remark that Smith's theorem only asserts that the union of Γ_i contains all the zeros of $P(z)$ and that any connected component of this consisting of just m disks Γ_i contains exactly m zeros. Therefore, we cannot conclude that each radius r_i gives the error bound for $x_i^{(11)}$, whereas we can say that each d_i does give. Finally we remark that $g(n,s) = g(5,\ 0.00066472493373) = 0.00267131066489 < 1$ so that we can also apply Zheng's result [15, Theorem A] to obtain

$$\| x^{(11)} - z^* \|_\infty \leqslant 0.000942577060002.$$

References

[1] O. Aberth, Iteration methods for finding all zeros of a polynomial simultaneously, *Math. Comp.* **27** (1973) 339–344.

[2] G. Alefeld, Componentwise inclusion and exclusion sets for solutions of quadratic equations in finite dimensional spaces, *Numer. Math.* **48** (1986) 391–416.

[3] E. Durand, *Solutions Numériques des Equations Algébriques, Tome I* (Masson, Paris, 1960).

[4] W.B. Gragg and R.A. Tapia, Optimal error bounds for the Newton–Kantorovich theorem, *SIAM J. Numer. Anal.* **11** (1974) 10–13.

[5] I.O. Kerner, Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, *Numer. Math.* **8** (1966) 290–294.

[6] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[7] J. Schröder, Nichtlineare Majoranten beim Verfahren der schrittweisen Näherung, *Arch. Math.* **7** (1956) 471–484.

[8] B.T. Smith, Error bounds for zeros of a polynomial based upon Gerschgorin's theorem, *J. Assoc. Comput. Mach.* **17** (1970) 661–674.

[9] M. Urabe, A posteriori componentwise error estimation of approximate solution to nonlinear equations, Lecture Notes in Comput. Sci. **29** (Springer, Berlin, 1975) 99–117.

[10] T. Yamamoto, An existence theorem for solution of nonlinear systems and its application to algebraic equations, *3rd USA–Japan Computer Conference Proc.* (AFIPS Press, Reston, VA and IPSJ, 1978) 300–304.

[11] T. Yamamoto, Componentwise error estimates for approximate solutions of nonlinear equations, *J. Inform. Process.* **2** (1979) 121–126.

[12] T. Yamamoto, Error bounds for computed eigenvalues and eigenvectors, *Numer. Math.* **34** (1980) 189–199.
[13] T. Yamamoto, Error bounds for computed eigenvalues and eigenvectors, II, *Numer. Math.* **40** (1982) 201–206.
[14] T. Yamamoto, Error bounds for approximate solutions of systems of equations, *Japan J. Appl. Math.* **1** (1984) 157–171.
[15] S. Zheng, On convergence of a parallel algorithm for finding the roots of a polynomial, *J. Math. Res. Exposition* **7** (1987) 657–660 (in Chinese).